
On the Rank of the Product of Certain Square Matrices*
and
Construction of Transformations to Canonical Forms†

By
WALTER O. MENGE



This dissertation is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the University of Michigan.

*Reprinted from *Bulletin of the American Mathematical Society*, February, 1932

†Reprinted from *American Journal of Mathematics*, Vol. LV, No. 4, October, 1933

On the Rank of the Product of Certain Square Matrices^{*}
and
Construction of Transformations to Canonical Forms[†]

By
WALTER O. MENGE



This dissertation is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the University of Michigan.

1932

^{*}Reprinted from *Bulletin of the American Mathematical Society*, February, 1932

[†]Reprinted from *American Journal of Mathematics*, Vol. LV, No. 4, October, 1933



ON THE RANK OF THE PRODUCT OF CERTAIN SQUARE
MATRICES

BY

W. O. MENGE

Reprinted from BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY
February, 1932

ON THE RANK OF THE PRODUCT OF CERTAIN SQUARE MATRICES*

BY W. O. MENGE

1. *Introduction.* This paper presents several theorems which were found during an investigation conducted by the author into the structure of matrices which transform given matrices into their so-called classic and rational canonical forms.† When the elementary divisors of a given matrix are known, these theorems completely determine the rank of a product of matrices of the form

$$\prod_{i=1}^{\omega} (A - \lambda_i I)^{k_i}.$$

An interesting proof of the Hamilton-Cayley theorem and a determination of the equation of minimum degree satisfied by a matrix are obtained from this point of view.

2. *Invariant Factors.* Consider the square matrix $A = (a_{ij})$ of order n with constant elements. If the n -rowed identity matrix be denoted by I , the characteristic matrix $(A - \lambda I)$ is defined as the matrix obtained by subtracting the variable λ from each principal diagonal element of A . The determinant, $D(\lambda)$, of the characteristic matrix $(A - \lambda I)$ is called the characteristic de-

* Presented to the Society, December 30, 1930. The author wishes to acknowledge his appreciation to J. A. Nyswander, University of Michigan, for many helpful suggestions throughout the progress of the work.

† Dickson, *Modern Algebraic Theory*, Chap. 5.

terminant of A . In order to secure complete generality in the subsequent discussion, the characteristic equation $D(\lambda) = 0$ is assumed to have the distinct roots $\lambda_1, \lambda_2, \dots, \lambda_\omega$ of multiplicities $n^{(1)}, n^{(2)}, \dots, n^{(\omega)}$, respectively. Hence

$$D(\lambda) = (\lambda - \lambda_1)^{n^{(1)}} (\lambda - \lambda_2)^{n^{(2)}} \dots (\lambda - \lambda_\omega)^{n^{(\omega)}},$$

where $\sum_{i=1}^{\omega} n^{(i)} = n$.

Designate by $G_j(\lambda)$ the highest common factor of all j th minors (of order $n-j$) of the characteristic determinant $D(\lambda)$. These common factors are chosen so that the coefficient of the highest power of λ is unity. Thus one can write

$$G_j(\lambda) = (\lambda - \lambda_1)^{m_j^{(1)}} (\lambda - \lambda_2)^{m_j^{(2)}} \dots (\lambda - \lambda_\omega)^{m_j^{(\omega)}}, \quad (j = 1, 2, \dots, n).$$

The invariant factors of the matrix A are defined as

$$D_j(\lambda) = G_{j-1}(\lambda)/G_j(\lambda) = (\lambda - \lambda_1)^{n_j^{(1)}} (\lambda - \lambda_2)^{n_j^{(2)}} \dots (\lambda - \lambda_\omega)^{n_j^{(\omega)}} \\ = \lambda^{n_j} - d_{1j}\lambda - d_{2j}\lambda^2 - \dots - d_{n_j,j}\lambda^{n_j-1}, \quad (j = 1, 2, \dots, n),$$

where

$$G_0(\lambda) = D(\lambda), \quad n_j^{(i)} = m_{j-1}^{(i)} - m_j^{(i)}, \quad n_j = \sum_{i=1}^{\omega} n_j^{(i)}, \\ (i = 1, 2, \dots, \omega; j = 1, 2, \dots, n).$$

3. *Classic Canonical Form.* Define the integer σ , so that

$$D_j(\lambda) \neq 1 \text{ for } j \leq \sigma; \quad D_j(\lambda) = 1 \text{ for } j > \sigma.$$

The classic canonical form for the n -rowed square matrix A is defined as

$$P = \begin{vmatrix} N_{11} & 0 & 0 & \dots & 0 \\ 0 & N_{21} & 0 & \dots & 0 \\ 0 & 0 & N_{31} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & N_{\omega\sigma} \end{vmatrix},$$

where N_{ij} is an $n_j^{(i)}$ -rowed square matrix of the form

$$N_{ij} = \begin{vmatrix} \lambda_i & 0 & 0 & \dots & 0 \\ 1 & \lambda_i & 0 & \dots & 0 \\ 0 & 1 & \lambda_i & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_i \end{vmatrix}, \quad \begin{matrix} (i = 1, 2, \dots, \omega; \\ j = 1, 2, \dots, \sigma). \end{matrix}$$

Each of the matrices N_{ij} is associated with a particular elementary divisor. From the above definition it can be easily verified that the invariant factors and elementary divisors of the matrix P are identical with the invariant factors and elementary divisors of the matrix A .

4. *Ranks.* It has been shown* that a necessary and sufficient condition that a non-singular matrix B exist such that

$$A = B \cdot P \cdot B^{-1}$$

is that the matrices A and P have the same invariant factors, or if we prefer, the same elementary divisors. It follows immediately that a non-singular matrix B exists such that

$$(1) \quad A - \lambda I = B \cdot (P - \lambda I) \cdot B^{-1}.$$

Since equation (1) is an identity in the variable λ , one can write

$$(2) \quad \prod_{i=1}^{\omega} (A - \lambda_i I)^{k_i} = B^m \cdot \prod_{i=1}^{\omega} (P - \lambda_i I)^{k_i} \cdot (B^{-1})^m \\ = B^m \cdot \prod_{i=1}^{\omega} (P - \lambda_i I)^{k_i} \cdot (B^m)^{-1},$$

where $m = \sum_{i=1}^{\omega} k_i$. Since the matrix B is non-singular it follows that B^m is also non-singular. Hence we may state the following lemma.

LEMMA 1. *The rank of the product*

$$\prod_{i=1}^{\omega} (A - \lambda_i I)^{k_i}$$

is identical with the rank of the product

$$\prod_{i=1}^{\omega} (P - \lambda_i I)^{k_i}.$$

It will be seen that the rank of any given product of the latter type can be determined by the invariant factors and elementary divisors of the matrix A .

Using the notation of Kronecker, define a matrix of order q , $(\delta_{r,s+1}) = (c_{rs})$, where $c_{rs} = 1$ for $r = s+1$, $c_{rs} = 0$ for $r \neq s+1$. It is easy to verify that $(\delta_{r,s+1})^k = (\delta_{r,s+k})$, and to prove the following result.

* Bôcher, *Introduction to Higher Algebra*, pp. 279-283.

LEMMA 2. *The matrix $(\delta_{r,s+k})$ is of rank $q-k$ if $q-k \geq 0$ and of zero rank if $q-k < 0$.*

In forming the product $(P - \lambda_i I)^{k_i}$ from its linear factors it is apparent from the special nature of each factor that the determinants $|N_{hj} - \lambda_i I|$ associated with each elementary divisor multiply each other independently of the determinants associated with other elementary divisors. The product is of the same form as each factor, and the determinant associated with each elementary divisor in the product is made up of the product of the corresponding determinants in the several factors. The determinants $|N_{ij} - \lambda_i I|$ are of the form $(\delta_{r,s+1})$ with $q = n_j^{(i)}$, and the determinants $|N_{hj} - \lambda_i I|$ with $h \neq i$ are non-singular.

Define $r_i^{(l)}$ as the number of elementary divisors of A divisible by $(\lambda - \lambda_i)^l$. Then $r_i^{(l)}$ will also be the number of elementary divisors of P which contain the factor $(\lambda - \lambda_i)^l$. Upon applying Lemma 2 for the multiplication of matrices of this form it is seen that the rank of the determinant associated with the elementary divisor $(\lambda - \lambda_i)^{n_j^{(i)}}$ in the product $(P - \lambda_i I)^{k_i}$ will be $n_j^{(i)} - k_i$ if $n_j^{(i)} - k_i \geq 0$ and will be zero if $n_j^{(i)} - k_i < 0$. The determinants associated with elementary divisors which do not contain the root λ_i will be non-singular. Thus the rank of the matrix of the product $(P - \lambda_i I)^{k_i}$ is seen to be

$$n - \sum_{l=1}^{k_i} r_i^{(l)}.$$

In the product $(P - \lambda_i I)^{k_i} (P - \lambda_j I)^{k_j}$, $j \neq i$, the determinants associated with elementary divisors involving either of the roots λ_i, λ_j are singular and of the form $(\delta_{r,s+1})$, while those associated with elementary divisors involving other roots are non-singular. Hence the rank of the product $(P - \lambda_i I)^{k_i} (P - \lambda_j I)^{k_j}$ is

$$n - \sum_{l=1}^{k_i} r_i^{(l)} - \sum_{l=1}^{k_j} r_j^{(l)}.$$

In general we find that the rank of the product

$$\prod_{i=1}^{\omega} (P - \lambda_i I)^{k_i}$$

is

$$n - \sum_{i=1}^{\omega} \sum_{l=1}^{k_i} r_i^{(l)}.$$

Upon applying Lemma 1 to the result just obtained we have the following theorem.

THEOREM 1. *The rank of the product*

$$\prod_{i=1}^{\omega} (A - \lambda_i I)^{k_i}$$

is

$$(i) \quad n - \sum_{i=1}^{\omega} \sum_{l=1}^{k_i} r_i^{(l)}.$$

This result may also be written in the form

$$(ii) \quad n - \sum_{i=1}^{\omega} \left[k_i r_i^{(k_i)} + \sum_{n_h^{(i)} < k_i} n_h^{(i)} \right].$$

Define $D_j(A)$ as the matrix obtained by replacing the variable λ by the matrix A in the j th invariant factor $D_j(\lambda)$. Thus

$$D_j(A) = (A - \lambda_1 I)^{n_j^{(1)}} (A - \lambda_2 I)^{n_j^{(2)}} \cdots (A - \lambda_{\omega} I)^{n_j^{(\omega)}}.$$

Any matrix $f(A)$, where $f(\lambda)$ is of equal or lower degree in λ than $D_j(\lambda)$ can be obtained from $D_j(A)$ in one of the following ways:

(a) by replacing one or more factors, say $(A - \lambda_1 I)$, by another, say $(A - \lambda_2 I)$;

(b) by omitting one or more factors, say $(A - \lambda_1 I)$, or by replacing them by non-singular matrices.

Under Case (a), the rank R of the altered matrix is, according to Theorem 1,

$$R = n - \sum_{l=1}^{n_j^{(1)}-1} r_1^{(l)} - \sum_{l=1}^{n_j^{(2)}+1} r_2^{(l)} - \sum_{i=3}^{\omega} \sum_{l=1}^{n_j^{(i)}} r_i^{(l)}.$$

Similarly the rank R_j of the matrix $D_j(A)$ is

$$R_j = n - \sum_{i=1}^{\omega} \sum_{l=1}^{n_j^{(i)}} r_i^{(l)},$$

and the difference in rank is

$$R - R_j = r_1^{(n_j^{(1)})} - r_2^{(n_j^{(2)}+1)}.$$

It is apparent, however, from the definition of $r_i^{(l)}$ and the prop-

erty of each invariant factor being divisible by each succeeding invariant factor, that

$$r_1^{(n_j^{(1)})} \geq j, \quad \text{while} \quad r_2^{(n_j^{(2)}+1)} < j.$$

From this it follows immediately that $R \geq R_j + 1$ and hence the replacement in Case (a) increases the rank by at least unity.

Under Case (b) by Theorem I the rank R of the altered matrix is determined as

$$R = n - \sum_{l=1}^{(n_j^{(1)}-1)} r_1^{(l)} - \sum_{i=2}^{\omega} \sum_{l=1}^{n_j^{(i)}} r_i^{(l)},$$

from which one concludes

$$R - R_j = r_1^{(n_j^{(1)})}.$$

Hence the rank R is larger than the rank R_j of the invariant factor matrix $D_j(A)$. Thus one has established the following fact.

THEOREM 2. *The rank of any invariant factor matrix $D_j(A)$ is the minimum for all matrices $f(A)$, where $f(\lambda)$ is of the same or lower degree in λ than $D_j(\lambda)$.*

By setting $k_i = n_j^{(i)}$, ($i=1, 2, \dots, \omega$), the result (ii) expressed in Theorem I can be applied to determine the rank R_j of $D_j(A)$. Here we find

$$(3) \quad R_j = n - \sum_{i=1}^{\omega} \left[n_j^{(i)} \cdot r_i^{(n_j^{(i)})} + \sum_{n_h^{(i)} < n_j^{(i)}} n_h^{(i)} \right],$$

the summation inside the bracket being with respect to h . Setting $s_j^{(i)} = r_i^{(n_j^{(i)})} - j$ in equation (3), we obtain

$$R_j = n - \sum_{i=1}^{\omega} \left[n_j^{(i)} (j + s_j^{(i)}) + \sum_{n_h^{(i)} < n_j^{(i)}} n_h^{(i)} \right],$$

and

$$(4) \quad R_j = n - j \cdot n_j - \sum_{i=1}^{\omega} \left[s_j^{(i)} n_j^{(i)} + \sum_{n_h^{(i)} < n_j^{(i)}} n_h^{(i)} \right],$$

since

$$\sum_{i=1}^{\omega} n_j^{(i)} = n_j.$$

Equation (4) may be written in the form

$$R_j = n - j \cdot n_j - \sum_{i=1}^{\omega} \sum_{h=j+1}^{\sigma} n_h^{(i)},$$

or

$$(5) \quad R_j = n - j \cdot n_j - \sum_{h=j+1}^{\sigma} n_h.$$

By virtue of the equation

$$\sum_{h=1}^j n_h = n - \sum_{h=j+1}^{\sigma} n_h,$$

the expression (5) may be written in the form

$$R_j = \sum_{h=1}^j n_h - j \cdot n_j = \sum_{h=1}^j (n_h - n_j).$$

THEOREM 3. *The rank R_j of $D_j(A)$ is*

$$\sum_{h=1}^j (n_h - n_j) = \sum_{h=1}^j n_h - j \cdot n_j = n - j \cdot n_j - \sum_{h=j+1}^{\sigma} n_h.$$

It should be noted that two well known properties of the matrix A follow directly from Theorem 3, namely, *Every square matrix satisfies its characteristic equation* and *The equation of minimum degree satisfied by the matrix A is the first invariant factor equated to zero.*

THEOREM 4. *In progressing from $D_{j+1}(A)$ to $D_j(A)$ by multiplying by successive linear factors, the rank decreases j for each linear factor.*

THEOREM 5. *The difference in the ranks of two consecutive invariant factor matrices, $D_{j+1}(A)$ and $D_j(A)$, is $j(n_j - n_{j+1})$.*

THE UNIVERSITY OF MICHIGAN

CONSTRUCTION OF TRANSFORMATIONS TO CANONICAL FORMS.

By WALTER O. MENGE.

1. *Introduction.* The problem of establishing the existence of a transformation to the rational canonical form has received considerable attention from various writers, each of whom has provided a characterization of the form. At least two of these have submitted existence proofs which are satisfactory from a modern standpoint. The first is the classical proof of Frobenius* and the second is due to Dickson and his predecessors and completed at an essential point by Bennett.† Each of these proofs employs only rational operations in the field of the coefficients of the original form and each is constructible in the sense that all the operations may be carried out in any special case in a finite number of steps. This paper therefore assumes the following:

THEOREM 1. *Let A be any n -rowed square matrix with elements in any field F . Then A is similar to a matrix P_r of the rational canonical form, that is, there exists a non-singular matrix B , with elements in F , such that*

$$B \cdot A \cdot B^{-1} = P_r.$$

Although the two proofs mentioned above afford constructions of the matrix B in a finite number of steps, the construction is extraordinarily laborious by either method. The purpose of this paper is to present a solution of the problem of finding a practicable method of determining the matrix B in the above theorem, given a particular A and hence a particular P_r .

In the latter part of the paper there is presented explicitly a transformation from the rational canonical form to the classic canonical form, and another for the reverse process. These forms for the transformation possess the advantage of directness and furnish simple practical schemes under entirely general hypotheses. Numerical examples are appended for the purpose of illustrating the method.

2. *Rational canonical form.* Consider the linear forms

$$(1) \quad X_i = \sum_{j=1}^n a_{ij}x_j \quad (i = 1, 2, \dots, n)$$

* Muth, *Elementarteiler* 3, and references given there.

† L. Dickson, "Modern algebraic theories," Chapter V; A. Bennett, *American Mathematical Monthly*, vol. 38, pp. 377-383.

in the independent variables $x_1, x_2, \dots, x_n$ with constant coefficients a_{ij} in any field F . The case in which the matrix A of the coefficients is singular is not excluded, so that the transformation (1) is entirely general.

Designate by $G_j(\lambda)$ the highest common factor of all j -th minors (of order $n - j$) of the characteristic determinant

$$(2) \quad D(\lambda) = |A - \lambda I|.$$

These common factors are chosen so that the coefficient of the highest power of the variable λ is unity. The invariant factors of the matrix A are defined as

$$(3) \quad \begin{aligned} D_j(\lambda) &= G_{j-1}(\lambda)/G_j(\lambda) \\ &= \lambda^{n_j} - d_{1j} - d_{2j}\lambda - \dots - d_{n_jj}\lambda^{n_j-1}, \quad (j = 1, 2, \dots, n). \end{aligned}$$

Define the integer σ , so that

$$D_j(\lambda) \neq 1 \text{ for } j \leq \sigma; \quad D_j(\lambda) = 1 \text{ for } j > \sigma.$$

It can be shown* that if transformation (1) be subjected to the introduction of new variables,

$$(4) \quad \begin{aligned} y_i &= \sum_{j=1}^n b_{ij}x_j & (i = 1, 2, \dots, n), \\ Y_i &= \sum_{j=1}^n b_{ij}X_j & (i = 1, 2, \dots, n), \end{aligned}$$

(the matrix B of the coefficients b_{ij} being non-singular) then the matrix P of the coefficients p_{ij} in the linear form

$$(5) \quad Y_i = \sum_{j=1}^n p_{ij}y_j \quad (i = 1, 2, \dots, n)$$

satisfies the equation

$$(6) \quad P = B A B^{-1}.$$

In the special case, when the linear form (5) simplifies to

$$(7) \quad \begin{aligned} Y_{1j} &= y_{2j} \\ Y_{2j} &= y_{3j} \\ &\vdots \\ Y_{n_j-1,j} &= y_{n_jj} \\ Y_{n_jj} &= \sum_{s=1}^{n_j} d_{sj}y_{sj} \end{aligned} \quad (j = 1, 2, \dots, \sigma),$$

the linear forms are said to be in the rational canonical form. This char-

* Dickson, *ibid.*

acterization is equivalent to that of Dickson. Each of the groups displayed in (7), corresponding to a particular invariant factor of the matrix A , is called a "chain." The coefficients d_{sj} in the last equation are precisely the coefficients of the several powers of the variable λ in the definition (3) of the invariant factors.

The matrix P_r of the coefficients of the rational canonical form (7) is

$$(8) \quad P_r = \left\| \begin{array}{cccccc} M_1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & M_2 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & M_3 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & M_\sigma \end{array} \right\|$$

where

$$(9) \quad M = \left\| \begin{array}{cccccc} 0 & 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ d_{1j} & d_{2j} & d_{3j} & d_{4j} & \cdot & \cdot & \cdot & d_{n_jj} \end{array} \right\| \quad (j = 1, 2, \cdot \cdot \cdot, \sigma).$$

Certain necessary conditions on the coefficients of the transformation (4) can be found by multiplying equation (6) on the right by B , yielding

$$(10) \quad B \cdot A = P_r \cdot B.$$

Upon equating corresponding elements in the two products comprising the members of this equation one obtains the following necessary conditions to be satisfied by the elements of the matrix B :

(for the first row)

$$(11) \quad \sum_{s=1}^n a_{sk} b_{1s} = b_{2k}, \quad (k = 1, 2, \cdot \cdot \cdot, n),$$

(for the second row)

$$(12) \quad \sum_{s=1}^n a_{sk} b_{2s} = b_{3k}, \quad (k = 1, 2, \cdot \cdot \cdot, n),$$

• • • • •

(for the $n_1 - 1$ -th row)

$$(13) \quad \sum_{s=1}^n a_{sk} b_{n_1-1,s} = b_{n_1k}, \quad (k = 1, 2, \cdot \cdot \cdot, n),$$

(for the n_1 -th row)

$$(14) \quad \sum_{s=1}^n a_{sk} b_{n_1s} = \sum_{s=1}^n d_{s1} b_{sk}, \quad (k = 1, 2, \cdot \cdot \cdot, n),$$

together with systems of equations similar in form to (11), (12), $\dots$, (13), and (14) corresponding to each of the other chains.

Let B_{i1} ($i = 1, 2, \dots, n_1$) denote respectively the one-rowed matrices formed by the first n_1 successive rows of the matrix B . Let B_{i2} ($i = 1, 2, \dots, n_2$) denote respectively the one-rowed matrices formed by the next n_2 successive rows of the matrix B , etc. Thus, in general

$$(15) \quad B_{ij} = (b_{s1}, b_{s2}, \dots, b_{sn})$$

where

$$s = \sum_{h=1}^{j-1} n_h + i.$$

Equations (11), (12), $\dots$ (13), and (14) can now be written in the simple form

$$(16) \quad \begin{array}{l} B_{11}A = B_{21} \\ B_{21}A = B_{31} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ B_{n_1-1,1}A = B_{n_1} \\ B_{n_1}A = \sum_{s=1}^{n_1} d_{s1}B_{s1}. \end{array}$$

The systems of equations for the succeeding chains corresponding to the second and succeeding invariant factors would appear as follows:

$$(17) \quad \begin{array}{l} B_{1j}A = B_{2j} \\ B_{2j}A = B_{3j} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ B_{n_{j-1},j}A = B_{n_jj} \\ B_{n_jj}A = \sum_{s=1}^{n_j} d_{sj}B_{sj}. \end{array} \quad (j = 2, 3, \dots, \sigma)$$

After the substitution of the values given in each equation of (16) and (17) into the succeeding equations in the same chain, one obtains

$$(18) \quad \begin{array}{l} B_{1j}A = B_{2j} \\ B_{1j}A^2 = B_{3j} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ B_{1j}A^{n_j-1} = B_{n_jj} \\ B_{1j} \cdot D_j(A) = 0. \end{array} \quad (j = 1, 2, \dots, \sigma),$$

where $D_j(A)$ is the matrix obtained by replacing the variable λ by the matrix A in the j -th invariant factor $D_j(\lambda)$.

When the notation B_{1j} is replaced by L_j ($j = 1, 2, \dots, \sigma$) it appears from equations (18) that the "leaders" L_j of the several chains must necessarily satisfy the several invariant factor equations

$$(19) \quad L_j \cdot D_j(A) = 0, \quad (j = 1, 2, \dots, \sigma).$$

Furthermore, by means of equations (18) every row of the matrix B is expressible in terms of the leader of the particular chain to which that row belongs. With the aid of equations (18) the successive rows of the matrix B can now be written sequentially

$$(20) \quad L_1 I, L_1 A, L_1 A^2, \dots, L_1 A^{n_1-1}, L_2 I, L_2 A, \dots, L_\sigma I, \dots, L_\sigma A^{n_\sigma-1}.$$

Since equations (11), (12), $\dots$, (13), and (14), and consequently equations (18) are necessary conditions on the elements of the matrix B it follows that every transformation to the rational canonical form is of the form (20). It remains to be shown how the leaders L_j ($j = 1, 2, \dots, \sigma$) can be so chosen as to produce a non-singular matrix B .

Denote the elements of the one-rowed matrix L_j by l_{ij} ($i = 1, 2, \dots, n$) so that

$$(21) \quad L_j = (l_{1j}, l_{2j}, \dots, l_{nj}) \quad (j = 1, 2, \dots, \sigma).$$

The matrix B can now be exhibited in the form:

$$(22) \quad B = \begin{vmatrix} & l_{11} & & l_{21} & \dots & & l_{n1} \\ \Sigma a_{i1}^{(1)} & l_{i1} & \Sigma a_{i2}^{(1)} & l_{i1} & \dots & \Sigma a_{in}^{(1)} & l_{i1} \\ \Sigma a_{i1}^{(2)} & l_{i1} & \Sigma a_{i2}^{(2)} & l_{i1} & \dots & \Sigma a_{in}^{(2)} & l_{i1} \\ & & & & & & \\ \Sigma a_{i1}^{(n_1-1)} & l_{i1} & \Sigma a_{i2}^{(n_1-1)} & l_{i1} & \dots & \Sigma a_{in}^{(n_1-1)} & l_{i1} \\ & l_{12} & & l_{22} & \dots & & l_{n2} \\ \Sigma a_{i1}^{(1)} & l_{i2} & \Sigma a_{i2}^{(1)} & l_{i2} & \dots & \Sigma a_{in}^{(1)} & l_{i2} \\ & & & & & & \\ \Sigma a_{i1}^{(n_\sigma-1)} & l_{i\sigma} & \Sigma a_{i2}^{(n_\sigma-1)} & l_{i\sigma} & \dots & \Sigma a_{in}^{(n_\sigma-1)} & l_{i\sigma} \end{vmatrix}$$

in which each summation extends from $i = 1$ to $i = n$ and $a_{ik}^{(h)}$ is the element in the i -th row and k -th column of the matrix A^h .

Each of the equations

$$(19) \quad L_j \cdot D_j(A) = 0 \quad (j = 1, 2, \dots, \sigma),$$

can be considered as a set of n linear homogeneous equations in the field F between the elements l_{ij} of the leader L_j . Since the rank of $D_1(A)$ is zero,* the elements of the first leader L_1 can be considered as independent. Denote the ranks of the matrices $D_j(A)$ ($j = 2, 3, \dots, \sigma$) by R_j ($< n$). Hence R_j of the elements in the leader L_j can be determined as linear homogeneous functions, with coefficients in the field F , of the remaining $n - R_j$ elements.

* Menge, "On the rank of the product, etc.," *Bulletin, American Mathematical Society*, Feb. 1932, pp. 88-94.

Designate the remaining $n - R_j$ elements by m_{ij} ($i = 1, 2, \dots, n - R_j$). Then each element of the leader L_j is expressible as a linear combination of the independent parameters m_{ij} . The determinant of the matrix B yields upon expansion a homogeneous polynomial of the n -th degree in the parameters m_{ij} . Since it has been shown that a non-singular matrix B exists, there will be at least one term in the expansion of the determinant with a non-zero coefficient. Suppose that such a term is

$$(20) \quad c \cdot m_{i_1 j_1}^{e_1} \cdot m_{i_2 j_2}^{e_2} \cdot \dots \cdot m_{i_k j_k}^{e_k}, \text{ where } \sum_{h=1}^k e_{i_h j_h} = n,$$

and c is a numerical coefficient not equal to zero. It should be noted that it may not be necessary to carry out the expansion of the determinant of B in order to find such a term. In most numerical examples this term can be found by inspection. Set

$$m_{i_2 j_2} = m_{i_3 j_3} = \dots = m_{i_k j_k} = 1$$

and all of the other parameters involved in the expansion of the determinant of B , with the single exception of $m_{i_1 j_1}$, equal to zero. After these substitutions the determinant of the matrix B simplifies to a polynomial, $p(m_{i_1 j_1})$, in the single parameter $m_{i_1 j_1}$. The non-singularity of the transformation B can now be assured by assigning to the parameter $m_{i_1 j_1}$ a value distinct from the several roots of the equation

$$p(m_{i_1 j_1}) = 0.$$

A numerical example illustrating the method is appended to this paper.

3. *Classic canonical form.* The second problem to be considered is the explicit representation of the transformation from the rational canonical form to the classic canonical form. Let the invariant factors of the matrix A , and hence also the invariant factors of the matrix P_r , be exhibited as

$$(21) \quad D_j(\lambda) = (\lambda - \lambda_1)^{n_j^{(1)}} (\lambda - \lambda_2)^{n_j^{(2)}} \dots (\lambda - \lambda_\omega)^{n_j^{(\omega)}}, \quad (j = 1, 2, \dots, \sigma).$$

When a set of linear forms simplifies to the form

$$(22) \quad \begin{aligned} Y_{1ij} &= \lambda_i y_{1ij} \\ Y_{2ij} &= \lambda_i y_{2ij} + y_{1ij} \\ Y_{3ij} &= \lambda_i y_{3ij} + y_{2ij} \\ &\vdots \\ Y_{n_j^{(i)} ij} &= \lambda_i y_{n_j^{(i)} ij} + y_{n_j^{(i)} - 1, ij} \\ (i &= 1, 2, \dots, \omega; \quad j = 1, 2, \dots, \sigma), \end{aligned}$$

the linear forms are said to be in the classic canonical form. As is seen from this definition, there is a chain similar in structure to that exhibited above

for every distinct root in every invariant factor of the matrix A . Exhibited as a matrix the coefficients of the classic canonical form (22) appear as

$$(23) \quad P_c = \left\| \begin{array}{cccccc} N_{11} & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & N_{21} & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & N_{n\sigma,\sigma} \end{array} \right\|,$$

where N_{ij} is an $n_j^{(i)}$ -rowed square matrix of the form

$$(24) \quad N_{ij} = \left\| \begin{array}{cccccc} \lambda_i & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 1 & \lambda_i & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & \lambda_i & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \lambda_i \end{array} \right\| \\ \cdot (i = 1, 2, \cdot \cdot \cdot, \omega; \quad j = 1, 2, \cdot \cdot \cdot, \sigma).$$

The classic canonical form can be considered from one viewpoint as a further reduction of the rational canonical form, in which each invariant factor chain is sub-divided into smaller chains corresponding to the several elementary divisors comprising the invariant factor. A natural extension of the method of the preceding section would involve the consideration of a method of constructing a non-singular matrix T which would transform the rational canonical form P_r into its classic canonical form P_c . More specifically, the problem at hand is the determination of a non-singular matrix T , such that

$$T \cdot P_r \cdot T^{-1} = P_c.$$

After multiplying this equation on the right by T , one has

$$(25) \quad T \cdot P_r = P_c \cdot T.$$

In order to effect simplicity in notation consideration at the outset will be limited to the part of the matrix T , denoted by T_j , which will convert the rational canonical form (7) corresponding to the j -th invariant factor into the classic canonical form defined by (22). The equation corresponding to (25) but referring only to the j -th invariant factor is

$$(26) \quad T_j \cdot M_j = Q_j \cdot T_j,$$

where

$$Q_j = \left\| \begin{array}{cccccc} N_{1j} & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & N_{2j} & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & N_{n_j j} \end{array} \right\|.$$

Define
$$\alpha_i = \sum_{k=1}^{i-1} n_j^{(k)} + 1, \quad (i = 1, 2, \dots, \omega).$$

By equating corresponding elements in the two products comprising the members of equation (26) one obtains the following necessary conditions to be satisfied by the elements t_{kl} of the matrix T_j . From the k -th row ($k = 1, 2, \dots, n_j$) one has

$$\left. \begin{aligned} (27) \quad & d_{1j} t_{a_i n_j} = \lambda_i t_{a_i 1} \\ (28) \quad & t_{a_i, l-1} + d_{lj} t_{a_i n_j} = \lambda_i t_{a_i l} \end{aligned} \right\} \quad (k = \alpha_i; \quad l = 2, 3, \dots, n_j),$$

and

$$(29) \quad d_{1j} t_{kn_j} = t_{k-1, 1} + \lambda_i t_{k1}$$

$$(30) \quad t_{k, l-1} + d_{lj} t_{kn_j} = t_{k-1, l} + \lambda_i t_{kl} \quad (\alpha_i < k < \alpha_{i+1}; \quad l = 2, 3, \dots, n_j).$$

For the rows in which $k = \alpha_i$ ($i = 1, 2, \dots, \omega$) equations (28) can be solved sequentially beginning with the last for the values of $t_{a_i l}$ ($l = 1, 2, \dots, n_j - 1$) in terms of $t_{a_i n_j}$. In this manner one finds

$$(31) \quad t_{a_i l} = g_l(\lambda_i) t_{a_i n_j} \quad (l = 1, 2, \dots, n_j; \quad i = 1, 2, \dots, \omega)$$

where

$$(32) \quad g_l(\lambda_i) = \lambda_i^{n_j - l} - \sum_{h=0}^{n_j - l - 1} d_{h+l+1} \lambda_i^h.$$

In deriving equations (31) the properties expressed by equations (27) were not employed, but it may be easily verified that these equations are satisfied identically in the elements $t_{a_i n_j}$. Upon replacing in equations (27) the elements $t_{a_i l}$ by their values as given by equations (31) one obtains

$$(33) \quad t_{a_i n_j} D_j(\lambda_i) = 0.$$

Equations (33) are identities in the elements $t_{a_i n_j}$ by virtue of the hypotheses regarding the roots of the invariant factors as given in (21). Thus equations (31) and (33) and hence also (27) and (28) are satisfied for any values of the elements $t_{a_i n_j}$ ($i = 1, 2, \dots, \omega$), and these elements may be regarded as parameters in terms of which all the elements in their rows of the matrix T_j can be expressed.

After solving in the same manner equations (30) for each element t_{kl} , one finds

$$(34) \quad t_{kl} = \sum_{p=0}^m \frac{1}{(m-p)!} g_l^{(m-p)}(\lambda_i) t_{\alpha_i + p, n_j},$$

where $\alpha_i < k < \alpha_{i+1}$; $m = k - \alpha_i$; $i = 1, 2, \dots, \omega$, and

$$g_l^{(n)}(\lambda_i) = \left[\frac{d^n g_l(\lambda)}{d\lambda^n} \right]_{\lambda=\lambda_i}.$$

Equations (29) were not employed in obtaining (34). However it may be shown that these equations are also satisfied identically in the elements t_{kn_j} by the expressions given in equations (34). Upon replacing the elements t_{kl} in equations (29) by the values given in equations (34) one finds after simplification

$$(35) \quad t_{kn_j} \cdot D_j(\lambda_1) + t_{k-1, n_j} \cdot D_j^{(1)}(\lambda_1) + \cdots + t_{a_i, n_j} \cdot D_j^{(m)}(\lambda_i) = 0.$$

Since λ_i is a root of the equation

$$D_j(\lambda) = 0$$

of multiplicity $n_j^{(i)} = \alpha_{i+1} - \alpha_i$, it follows immediately that the coefficients of t_{hn_j} ($h = \alpha_i, \alpha_i + 1, \cdots, k$) in equations (35) vanish. Hence (34) and (35) and hence (29) and (30) are satisfied for any values of the elements t_{hn_j} ($h = 1, 2, \cdots, n_j$) and these elements may be considered as parameters in terms of which all other elements in the same row of T_j can be expressed.

Since equations (34) for the particular values $k = \alpha_i$, ($i = 1, 2, \cdots, \omega$), reproduce equations (32), one can now write the matrix T_j in the form

$$T_j = (t_{kl}) = \left(\sum_{p=0}^m \frac{1}{(m-p)!} g_i^{(m-p)}(\lambda_i) t_{a_i+p, n_j} \right)$$

where

$$m = k - \alpha_i; \quad \alpha_i < k < \alpha_{i+1},$$

in which the elements t_{kn_j} are to be considered as parameters and may have any values which insure the non-singularity of the matrix T_j .

Let

$$(36) \quad \begin{cases} t_{kn_j} = 1, & \text{for } k = \alpha_i; i = 1, 2, \cdots, \omega, \\ t_{kn_j} = 0, & \text{for } k \neq \alpha_i; i = 1, 2, \cdots, \omega. \end{cases}$$

For this particular choice of the parameters the matrix T_j takes the special form

$$(37) \quad T_j = (t_{kl}) = \left(\frac{1}{m!} g_i^{(m)}(\lambda_i) \right)$$

where $m = k - \alpha_i$, $\alpha_i < k < \alpha_{i+1}$; $i = 1, 2, \cdots, \omega$.

It remains to be shown that T_j as expressed in form (37) is non-singular. After the functions $g_i^{(m)}(\lambda_i)$ have been replaced by polynomials in λ_i according to the definitions (32) each element of the l -th column ($l = 1, 2, \cdots, n_j$) of the determinant of (37) is expressed as the sum of $n_j - l + 1$ terms. The determinant of (37) can then be expanded as the sum of $n_j(n_j + 1)/2$ determinants, each element of each determinant involving one term taken from the corresponding element of the determinant of (37). After removing common factors it is apparent that the new determinants found in

this manner are all zero (by virtue of possessing two identical columns) except the single determinant

$$(38) \quad \left| \frac{(\lambda_i^{n_i-l})^{(m)}}{m!} \right|.$$

The determinant (38) has been shown* to be non-singular and hence the matrix T_j as expressed explicitly by (37) is non-singular.

It is obvious that the transformation,

$$(39) \quad T = \left\| \begin{array}{cccccc} T_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & T_2 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & T_\sigma \end{array} \right\|,$$

in which each submatrix T_j is of the form (37), is non-singular and satisfies the matrix equation

$$T \cdot P_r \cdot T^{-1} = P_c.$$

It follows immediately that the matrix $C = T \cdot B$ is non-singular and satisfies the matrix equation

$$C \cdot A \cdot C^{-1} = P_c,$$

and that the matrix $S = T^{-1}$ satisfies the matrix equation

$$S \cdot P_c \cdot S^{-1} = P_r.$$

A numerical example of the method of this section is appended.

4. *Numerical illustration.* In order to illustrate the method of the previous section a numerical example of the transformations previously derived in general form, will not be presented. Let the given linear form † be

$$(A) \quad \begin{aligned} X_1 &= -2x_1 - x_2 - x_3 + 3x_4 + 2x_5 \\ X_2 &= -4x_1 + x_2 - x_3 + 3x_4 + 2x_5 \\ X_3 &= x_1 + x_2 - 3x_4 - 2x_5 \\ X_4 &= -4x_1 - 2x_2 - x_3 + 5x_4 + x_5 \\ X_5 &= 4x_1 + x_2 + x_3 - 3x_4. \end{aligned}$$

After evaluating the characteristic determinant of the coefficients in the linear forms above, one finds

$$|A - \lambda I| = D(\lambda) = (\lambda - 2)^3 (\lambda + 1)^2.$$

* Nyswander, "A direct solution of systems of linear differential equations, etc," *American Journal of Mathematics*, vol. 47 (1925), pp. 272-3.

† Burnside, *Proceedings of the London Mathematical Society*, vol. 30 (1899), pp. 191-2.

The ranks of the determinants $|A - 2I|$ and $|A + I|$ are 3 and 4, respectively. It follows immediately* that the invariant factors of the matrix A must be, respectively,

$$D_1(\lambda) = (\lambda - 2)^2(\lambda + 1)^2 = \lambda^4 + 4\lambda^3 - 3\lambda^2 - 2\lambda, \quad D_2(\lambda) = \lambda - 2.$$

The matrix equation,

$$L_2 D_2(A) = 0,$$

to be satisfied by the leader of the second chain, can now be written in the form (dropping the second subscript)

$$\begin{aligned} -4l_1 - 4l_2 + l_3 - 4l_4 + 4l_5 &= 0 \\ -l_1 - l_2 + l_3 - 2l_4 + l_5 &= 0 \\ -l_1 - l_2 - 2l_3 - l_4 + l_5 &= 0 \\ +3l_1 + 3l_2 - 3l_3 + 3l_4 - 3l_5 &= 0 \\ +2l_1 + 2l_2 - 2l_3 + l_4 - 2l_5 &= 0. \end{aligned}$$

These simultaneous equations have the solution

$$l_1 = p_{12}; \quad l_2 = p_{22}; \quad l_3 = 0; \quad l_4 = 0; \quad l_5 = p_{12} + p_{22},$$

where p_{12} and p_{22} are arbitrary parameters.

The transpose (written for convenience) of the matrix B can now be exhibited in the form

$$\begin{aligned} \left| \begin{array}{ccccc} p_{11}, & -2p_{11} - 4p_{21} + p_{31} - 4p_{41} + 4p_{51}, & 3p_{11} - p_{21} - 2p_{31} - p_{41} + p_{51}, & & \\ p_{21}, & -p_{11} + p_{21} + p_{31} - 2p_{41} + p_{51}, & -4p_{11} & + 4p_{31} - 8p_{41} + 4p_{51}, & \\ p_{31}, & -p_{11} - p_{21} & - p_{41} + p_{51}, & 2p_{11} + 2p_{21} - p_{31} + 2p_{41} - 2p_{51}, & \\ p_{41}, & 3p_{11} + 3p_{21} - 3p_{31} + 5p_{41} - 3p_{51}, & 3p_{11} + 3p_{21} - 3p_{31} + 7p_{41} - 3p_{51}, & & \\ p_{51}, & 2p_{11} + 2p_{21} - 2p_{31} + p_{41} & - p_{11} - p_{21} + p_{31} - 5p_{41} + 5p_{51}, & & \\ & & - 4p_{11} - 12p_{21} + 3p_{31} - 12p_{41} + 12p_{51}, & p_{12} & \\ & & - 12p_{11} - 4p_{21} + 12p_{31} - 24p_{41} + 12p_{51}, & p_{22} & \\ & & - 3p_{11} - 3p_{21} + 2p_{31} - 3p_{41} + 3p_{51}, & 0 & \\ & & 9p_{11} + 9p_{21} - 9p_{31} + 17p_{41} - 9p_{51}, & 0 & \\ & & - 3p_{11} - 3p_{21} + 3p_{31} - 15p_{41} + 11p_{51}, & p_{12} + p_{22} & \end{array} \right| \end{aligned}$$

in which all of the parameters p_{ij} are arbitrary, subject only to the condition that the matrix is non-singular. The coefficient of the term $p_{11}^4 \cdot p_{12}$ in the determinant of the above matrix is 81 and hence the matrix B will be non-singular if $p_{11} = p_{12} = 1$, and $p_{21} = p_{31} = p_{41} = p_{51} = p_{22} = 0$. After these substitutions the matrix B becomes

* Menge, *ibid.*

$$(B) = \left\| \begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ -2 & -1 & -1 & 3 & 2 \\ 3 & -4 & 2 & 3 & -1 \\ -4 & -12 & -3 & 9 & -3 \\ 1 & 0 & 0 & 0 & 1 \end{array} \right\|, \text{ and } B \cdot A \cdot B^{-1} = \left\| \begin{array}{ccccc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -4 & -4 & 3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{array} \right\|,$$

the last matrix being the required rational canonical form of A .

Proceeding to the problem of determining the transformation T from the rational canonical form to the classic canonical form, one finds

$$D_1(\lambda) = \lambda^4 + 4 + 4\lambda - 3\lambda^2 - 2\lambda^3$$

$$g_1(\lambda) = \lambda^3 + 4 - 3\lambda - 2\lambda^2$$

$$g_2(\lambda) = \lambda^2 - 3 - 2\lambda$$

$$g_3(\lambda) = \lambda - 2$$

$$g_4(\lambda) = 1.$$

After evaluating these polynomials and their first derivatives for $\lambda = 2$ and -1 , respectively, it is possible to write down immediately the transformation T ,

$$(T) = \left\| \begin{array}{cccc|c} -2 & -3 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ \hline 4 & 0 & -3 & 1 & 0 \\ 4 & -4 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array} \right\|, \text{ and } T \cdot P_r \cdot T^{-1} = \left\| \begin{array}{cccc} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 2 \end{array} \right\| = (P_c).^*$$

UNIVERSITY OF MICHIGAN.

* Added in reading proof: There has been recently published another proof of the theorem mentioned in the introduction, see Ingraham, *Bulletin of the American Mathematical Society*, vol. 39 (1933), pp. 379-382.

